

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 28, 2003, 18:27:17 ; Search time 31.3102 Seconds
(without alignments)
107.116 Million cell updates/sec

Title: US-09-743-225-10

Sequence: 1 CATTLRVYKGGGXA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23:*

1: sp_archaea;*

2: sp_bacteria;*

3: sp_fungi;*

4: sp_human;*

5: sp_invertebrate;*

6: sp_mammal;*

7: sp_mhc;*

8: sp_organelle;*

9: sp_phage;*

10: sp_plant;*

11: sp_rabbit;*

12: sp_virus;*

13: sp_vertebrate;*

14: sp_unclassified;*

15: sp_rvirus;*

16: sp_bacteriophage;*

17: sp_archeap;*

Q9awt7 oryza sativ
Q9sym0 arabidopsis
Q8lbn5 arabidopsis
Q9ya1 aeropyrum p
Q0083 saccharomyces
Q8yt58 arabidopsis
Q8uy4 streptococcus
Q8ukn5 agrobacteri
Q8ts14 methanobacter
Q9fn06 arabidopsis
Q943j5 oryza sativ
Q8bj16 mus musculus
Q8j943 fugu rubrip
Q8bq16 mus musculus
P74210 synchrocyt
Q9922 homo sapien
Q9k41 clostridium
Q8h7z2 oryza sativ
Q8kt00 candidatus
Q964n0 drosophila
Q9bj13 ornithodoros
Q8ny08 ornithodoros
Q25227 helicobacte
Q9z1z1 helicobacte
Q8ugs4 agrobacteri
Q9kef2 bacillus ha
Q9nl20 oryzolagrus
Q97yy8 sulfolobus
Q1592 caenorhabdi
Q31002 vibrio angu

ALIGNMENTS

RESULT 1

Q9u189 PRELIMINARY; PRT; 605 AA.
ID Q9u189
AC Q9u189;
DT 01-MAY-2000 (TREMBLrel 13, Created)
DT 01-MAY-2000 (TREMBLrel 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel 23, Last annotation update)
DE zeta tubulin (Gamma tubulin-related protein).
GN L338.0.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatida; Leishmania.
RN [1]NCI_TAXID=564;
RN SEQUENCE FROM N.A.
RP STRAIN=Friedlin;
RC STRAIN=Friedlin;
RA Weiler H., Hilbert H., Duesterhoeft A., Ivens A.C., Murphy L., Quail M., Rajandream M.A., Barrell B.G.; Submitted (RCV-2000) to the EMBL/GenBank/DBJ databases.
RN (2)SUBMISSIONS
RN SEQUENCE FROM N.A.
RP STRAIN=Friedlin;
RX MEDLINE=9814435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M., Smith D.F., "A physical map of the Leishmania major Friedlin genome.", Genome Res. 8:135-145(1998); DR InterPro: IP000217; Tubulin.
RA All13468; CAB63128.1; DR InterPro: IP000886; ER target.
RA DR InterPro: IP000217; Tubulin.
RA DR InterPro: IP003008; Tubulin_FtsZ.
RA DR InterPro: IP004058; Zeta_tubulin.
RA DR Pfam: PF00091; tubulin; 1.
RA DR PRINTS: PRO1516; TUBULIN.
RA DR PROSITE: PS00014; ER_TARGET; 1.
RA DR GTP-binding.
KW SQ SEQUENCE 605 AA; 65700 MW; 49944A374F5D8B82 CRC64;
Query Match 63.6%; Score 42; DB 5; Length 605;

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	63.6	605	5 Q9u189	Q9u189 leishmania s
2	41	62.1	569	16 Q8xvh4	Q8xvh4 ralstonia s
3	40	60.6	276	16 Q8d4t7	Q8d4t7 vibrio vuln
4	40	60.6	381	16 Q8db50	Q8db50 vibrio vuln
5	40	60.6	384	16 Q8zav6	Q8zav6 salmonella
6	40	60.6	607	16 Q92ur7	Q92ur7 rhizobium m
7	40	60.6	619	16 Q9ksf7	Q9ksf7 vibrio chol
8	40	60.6	932	10 Q9lxL3	Q9xL3 arabidopsis
9	40	60.6	937	10 Q8lgU3	Q8lgU3 arabidopsis
10	40	60.6	937	10 Q8L5J2	Q8L5J2 arabidopsis
11	40	60.6	938	10 Q8lnw2	Q8lnw2 arabidopsis
12	39	59.1	173	5 Q8IT03	Q8lu3 caenorhabdi
13	39	59.1	241	5 Q8MR7	Q8MR7 caenorhabdi
14	39	59.1	285	16 Q98em3	Q98em3 rhizobium 1
15	39	59.1	456	5 Q8To74	Q8To74 apis mellif
16	39	59.1	631	16 Q808E6	Q808E6 agrobacteri

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Best Local Similarity 54.5%; Pred. No. 23; Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;	Query Match Score 40; DB 16; Length 276; Best Local Similarity 77.8%; Pred. No. 24; Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
YY 1 CATTLYRKGGG 11 Db 383 CTALARTHEGGG 393	Qy 3 TLRVYKGGG 11 Db 193 TTRVYRGGG 201
RESULT 2	RESULT 4
Q8XVH4 PRELIMINARY; PRT; 569 AA. ID Q8XVH4 AC Q8XVH4 DT 01-MAR-2002 (TREMBLrel. 20, Created) DT 01-MAR-2002 (TREMBLrel. 21, Last sequence update) DT 01-MAR-2002 (TREMBLrel. 21, Last annotation update) DE Probable long-chain fatty-acid-CoA ligase protein (EC 6.2.1.3). RA FADDI OR RSC285 OR RS0048. DS Ralstonia solanacearum (Pseudomonas solanacearum). DC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; DC Ralstoniaceae; Ralstonia. DN NCBI_TAXID=305; RN [1] RP SEQUENCE FROM N.A. STRAIN=GMI0000; MEDLINE=21561879; PubMed=11823852; XX Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S., Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L., Chandler M., Choisne N., Claudelet-Renard C., Cunnac S., Demattei N., Gaspin C., Lavia M., Moisan A., Robert C., Saurin W., Schiex T., Siguer P., Thebaud P., Whalen M., Wincker P., Levy M., Weissbach J., Boucher C.A.; "Genome sequence of the plant pathogen Ralstonia solanacearum."; Nature 415:497-502 (2002). EMBL; AL646072; CAD16564.1; InterPro; IPR000873; AMP-bind. Pfam; PF000501; AMP-binding. PROSITE; PS00455; AMP_BINDING_1. DR LIGASE; Complete proteome. RN SEQUENCE: 569 AA; 62332 MW; F9D8556300638F54 CRC64; RN	Q8DESO PRELIMINARY; PRT; 381 AA. ID Q8DESO AC Q8DESO DT 01-MAR-2003 (TREMBLrel. 23, Created) DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update) DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update) DE Conserved hypothetical protein. GN VV10515 OS Vibrio vulnificus. OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; OC Vibrionaceae; Vibrio. OX NCBI_TAXID=672; RN [1] RP SEQUENCE FROM N.A. STRAIN=CMCP6; XX Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H., Choy H.E.; "Complete genome sequence of Vibrio vulnificus CMCP6."; Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases. EMBL; AE016812; AAC08099.1; RN
RESULT 3	RESULT 5
Q8D4T7 PRELIMINARY; PRT; 276 AA. ID Q8D4T7 AC Q8D4T7 DT 01-MAR-2003 (TREMBLrel. 23, Created) DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update) DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update) DE Uncharacterized protein conserved in bacteria. DN VV1202. DS Vibrio vulnificus. DC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; DC Vibrionaceae; Vibrio. DN NCBI_TAXID=672; RN [1] RP SEQUENCE FROM N.A. STRAIN=CMCP6; XX Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H., Choy H.E.; "Complete genome sequence of Vibrio vulnificus CMCP6."; Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases. EMBL; AE016812; AAC08099.1; RN	Q8Z8V6 PRELIMINARY; PRT; 384 AA. ID Q8Z8V6 AC Q8Z8V6 DT 01-MAR-2002 (TREMBLrel. 20, Created) DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update) DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update) DE Hypothetical protein Stry0480. GN STY0480 OS Salmonella typhi. OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; OC Enterobacteriaceae; Salmonella. OX NCBI_TAXID=601; RN [1] RP SEQUENCE FROM N.A. STRAIN=NCCT18; XX MEDLINE=21534947; PubMed=11677608; RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebasthia M., Baker S., Basham D., Brooks K., Chillingworth T., Connerton P., Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J., Feltham T., Hamlin N., Hague A., Hien T.T., Holroyd S., Jagels K., Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C., Quail M., Rutherford K., Whitehead S., Barrell B.G.; "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi C18."; Nature 413:848-852 (2001). EMBL; AL622766; CAD08897.1; DR InterPro; IPR06597; Sel-like. DR SMART; SM0671; SEV1.4; RN

Waterson N.; Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium. *Science* 282:2012-2018 (1998).

[2] RPR SEQUENCE FROM N.A.
STRAIN-BRISTOL N2;

[3] RCR GRAVES T.;
RA "The sequence of *C. elegans* cosmid Y73B6BL.";
RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.

[4] RPR SEQUENCE FROM N.A.
STRAIN-BRISTOL N2;

Waterson N.; Submitted (Nov-2002) to the EMBL/GenBank/DBJ databases.

[5] RCR EMBL: AC084197; AACN63427.1; -
RA DRD HYPOTHETICAL PROTEIN;
RL SEQUENCE 173 AA; 18222 MW; 01F73163B024C59E2 CRC64;
RN SQW

Waterston R.; RA RT "Genome sequence of the nematode *C. elegans*: a platform for investigating biology." The *C. elegans* Sequencing Consortium. *Nature* 409: 172-181.

science 282:2012-2018 (1998).
[2] SEQUENCE FROM N.A.;
STRAIN=BRLSTO1_N2;
Graves T.;
"The sequence of *C. elegans* cosmid Y73B6BL.";
Submitted (Oct 2000) to the ENSEMBL/GenBank/DBJ databases.

131
 SEQUENCE FROM N.A.
 STRAIN=Bristol N2;
 Waterston R.;
 Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 EMBL: AC084197; AACM4399.1;
 InterPro: IPR005818; Histone_H1/H5;
 InterPro: IPR005819; Histone_H5;
 InterPro: IPR003216; Linkerhist_N.
 Pfam: PF00538; linker_histone; 1.
 PRINTS: PR00624; HISTONEH5.
 ProdDom: PDD000373; Linkerhist_N; 1.
 DR

DR smali; savojo; nju; 1.
KW Hypothetical Protein.
SQ SEQUENCE 241 AA; 25409 MW;
5E73BDDDEEE1A692 CRC64;

	Matches	7; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
Y	5	RVYRGGG 11				
b	111111	RVYRGGG 34				

T Mesorhizobium loti; ";
T DNA Res. 7:331-338 (2000).
L EMBL: AP003003; BAB50895.1; -.
R InterPro: IPR004165; CoA_trans.
R Pfam: PF01144; CoA_trans; 1.
W Transferase; Complete proteome.
W SEQUENCE 285 AA; 30987 MW;
Q 2E611B62106EED CRG64;

RESULT 15
8BT04
D 08T04
C Q8T04;
PRELIMINARY;
PRT; 456 AA.

02-JUN-2002 (TRMBLrel. 21, Last sequence update)
 01-MAR-2003 (TRMBLrel. 23, Last annotation update)
 Dopamine receptor type D2.

Apis mellifera (Honeybee); Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis; NCBI_TAXID:7460;

SEQUENCE FROM N.A.
TISSUE-Brain, muscle;
MEDLINE-#98194783; PubMed-#5335160;
Ebert P.R., Rowland J.E., Toma D.P.;
"Isolation of seven unique biogenic amine receptor clones from the
honey bee by library scanning";

Insect Mol. Biol. /131-162(1998).
 [2] SEQUENCE FROM N.A.
 RP TISSUE-Brain mushroom bodies;
 RC Hirudinea M a Mustard J Hunter S J Mercer A Ward V

RA Ebert, P.R.;
 RT "An invertebrate D2 type dopamine receptor exhibits plasticity of expression in the mushroom bodies of the honey bee brain correlated with behavioural maturation of adults.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AF498306; AAM19330_1;
 DR InterPro: IPR00276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7em_1; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 RW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 456 AA; 52309 MW: 4486335470B367CA CRC64;
 Query Match Score 39; DB 5; Length 456;
 Best Local Similarity 66.7%; Pred. No. 63;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 3 TLRVYKGG 11
 |||:|||:
 Db 257 TLRVYKGG 265

Search completed: August 28, 2003, 18:38:02
 Job time : 33.3182 secs